Available online at www.sciencedirect.com

ScienceDirect

Energy Procedia 40 (2013) 95 – 101

Energy Procedia

European Geosciences Union General Assembly 2013, EGU

Division Energy, Resources & the Environment, ERE

Wind Forecasting using HARMONIE with Bayes Model Averaging for Fine-Tuning

Martin B. Peters, Enda O'Brien*, Alastair McKinstry, Adam Ralph

Irish Centre for High-End Computing, Trinity Technology &Enterprise Campus, Grand Canal Quay, Dublin 2, Ireland

Abstract

Wind-speed forecasts for a wind-farm in southwest Ireland were made for over one year using the operational HARMONIE mesoscale weather forecast model, and Bayes Model Averaging (BMA) for statistical post-processing to remove systematic local bias. The deterministic forecasts alone generated mean absolute errors of 1.7−2.0 ms^{-1} out to 24hrs, when interpolated to the location of the met-mast. Application of BMA reduced these errors by about 15%, to 1.5−1.6 ms^{-1}, on average. Forecast errors do not degrade significantly as forecast lead-time increases, at least out to 24 hours.

Keywords: wind-farm; forecasting; HARMONIE; Bayes Model Averaging

1. Introduction

At least two distinct elements are required to make accurate wind-speed forecasts for wind-farms: first, deterministic output from a weather forecast model, interpolated to the wind-farm site; and second, probabilistic or statistical post-processing to account for local biases or systematic errors in the model.

This paper reports on the skill achieved in forecasting wind-speeds for a wind-farm in the southwest of Ireland for over one year, using the HARMONIE mesoscale weather forecast model [1] along with Bayes

* Corresponding author. Tel.: +353-91-495946
E-mail address: enda.obrien@ichec.ie

doi:10.1016/j.egypro.2013.08.012

Model Averaging (BMA) [2] for statistical post-processing. BMA allowed forecasts to be made in a calibrated probabilistic format.

The main metric used to evaluate the forecasts is "Mean Absolute Error" (MAE). If f_i is the predicted value at time i, and y_i is the actual (observed) value, then MAE is defined as:

$$MAE = (1/n) \sum_{i=1}^{n} |f_i - y_i| \tag{1}$$

The questions addressed here include: how does MAE depend on the configuration of the HARMONIE weather model, and on the forecast lead-time? How much added value is provided by BMA, in reducing MAE? How does MAE depend on geographic location of the station of interest, and on the time of year? Other relevant questions are more values-based and not addressed here, such as, how small does MAE have to be to make the forecasts worthwhile?

2. Data and Forecast Models

Verifying wind-speed and other meteorological observations from the met mast of the wind farm were available for November 2010, and for all of 2012 (with some gaps). These data were measured at 64m above the surface (at turbine height). Three different variants of HARMONIE (as listed in Table 1) were used to make 24-hr forecasts during November 2010. Output from the operational version of HARMONIE (with 2.5km horizontal grid resolution), run 4 times per day by Met Éireann (the Irish meteorological service) was used for the forecasts during 2012. Finally, an experimental high-resolution version of HARMONIE (with 1km horizontal grid resolution) was used to make forecasts for a 30-day period spanning October-Nov. 2012.

3. Bayes Model Averaging

Bayes Model Averaging (BMA) is a statistical ensemble post-processing method that produces probabilistic forecasting based on the generated predictive probability density functions (PDFs) [3]. When dealing with model uncertainty, it often occurs that several models fit the data almost equally well but make different predictions for the same variable of interest. The BMA predictive PDF for a future quantity (e.g., wind-speed) is a weighted sum of the individual model PDFs, where the weights are calculated from the observations and individual model forecasts during a "training period" (typically 20-30 days for present purposes). The better component models typical obtain larger weighting than the less accurate models, but the key feature is that the BMA prediction based on *all* models is generally better than that based on any single model by itself. Even the less-accurate component models may contain useful forecast information that is integrated by BMA into an overall predictive PDF.

4. Results

4.1. Forecasts with Ensemble of Three HARMONIE variants for Nov. 2010

Table 1 lists the three different configurations of HARMONIE that were used to forecast wind-speeds during Nov. 2010. All three variants had a horizontal grid resolution of 2.5km. The "small domain" consisted of 300x300 horizontal points, while the "large domain" had 540x500 points. All models had 60 vertical levels. The "Alaro" physics package is designed for relatively coarse-resolution models (5km or larger grid-spacing), while the "Arome" package is designed primarily for higher-resolution models. Of the three configurations, the third one (non-hydrostatic, using Arome physics, on a large domain) is the most advanced from a physical and numerical point of view, but also the most computationally expensive.

All three models were run once per day for 24hr forecasts, starting at 00Z. Forecast quality usually degrades gradually as lead-time increases, so forecasts at hour 23 are typically worse than at hour 1. However, such degradation was negligibly small for the forecasts shown here.

Figure 1 shows the full 30-day time-series during November 2010 (at 1-hour intervals) for the observed wind-speed (black curve) and each of the three forecast models (coloured curves) interpolated to the met-mast location and height of 64m. Each model successfully forecasts the gross pattern of variability observed, though it appears they have a slight negative bias overall.

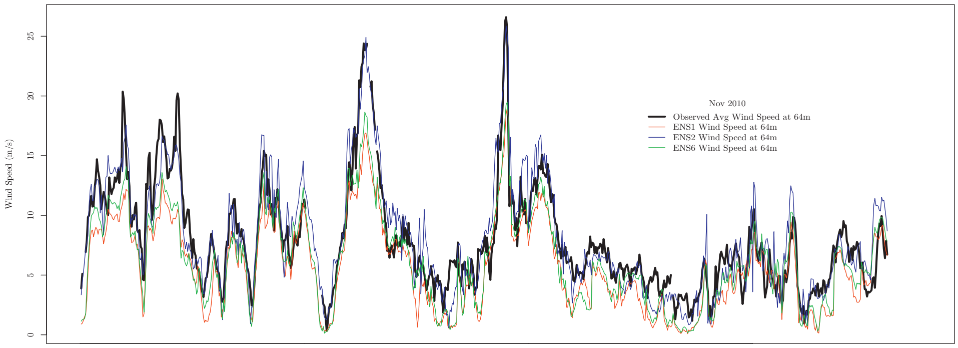

Fig. 1. 30-day time-series of HARMONIE wind-speed forecasts (coloured curves) vs. observations (black curve).

Table 1 distils the time-series of Fig. 1 into the numerical skill metrics (squared-correlation R^2, and MAE) achieved by each of the HARMONIE configurations. Not surprisingly, the more advanced "Arome" variant has a better correlation with observations (R^2=0.82) and a lower MAE (1.50 m s^{-1}) than the other two, over this 30-day period. With a mean wind-speed in Fig. 1 of approx. 9 m s^{-1}, a MAE of 1.5 m s^{-1} represents an error of 17%. However, this percentage error is amplified by periods of weak winds, such as the last 10 days in Fig. 1.

To apply BMA analysis to this small ensemble of HARMONIE forecasts, the first 26 days of the month were taken for "training" purposes. The BMA model was then used to forecast wind-speed for each hour of the next 4 days. When compared against the verifying observations for these 4 days, the BMA model had a MAE of 1.52 m s^{-1}; the "Arome" version had an MAE of 1.51 m s^{-1}, while the 3-member ensemble together had an MAE of 1.62 m s^{-1}. So in this case BMA was unable to add any real value from the two "Alaro" variants to the more skilful "Arome" version.

Table 1 Different HARMONIE variants used to forecast winds during Nov. 2010, and their skill metrics for wind-speed forecasts (at turbine height of 64m) vs. observations. R^2 is the squared correlation coefficient; MAE is Mean Absolute Error.

Model	64m R^2	MAE (m s^{-1})
Alaro physics, Hydrostatic, small domain	0.79	2.41
Alaro physics, Non-hydrostatic, small domain	0.79	2.04
Arome physics, Non-hydrostatic, large domain	0.82	1.50

Two useful and related BMA products are "threshold forecasting" (as shown in Fig. 2) and "quartile forecasting" (as shown in Fig. 3). Threshold forecasting provides probabilities that wind-speed will be below a given threshold (e.g., 8 m s^{-1} in Fig. 2). The black curve in Fig. 2 shows observed hourly winds for the 4 days being forecast by BMA, while the blue curve shows the probability of the wind speed being below 8 m s^{-1}, at each forecast hour. (Probabilities are scaled by a factor of 10 to match the wind-speed scale on the y-axis). Thus, high values of the blue curve typically correspond to low winds, and vice versa.

Quartile forecasting predicts wind speed strengths associated with specified probabilistic quartile values. Thus in Fig. 3, BMA predicts a 25% probability that wind-speed will be lower than the red curve value for each forecast hour, and a 75% probability that wind-speed will be lower than the blue curve. The green curve shows the 50% probability of wind-speed being lower than this value – in other words it shows the most likely or expected wind-speed. The black curve shows the verifying observed values. Not surprisingly, perhaps, it is outside the "envelope" bounded by the 25%-75% probability curves somewhere between 25% and 50% of the time.

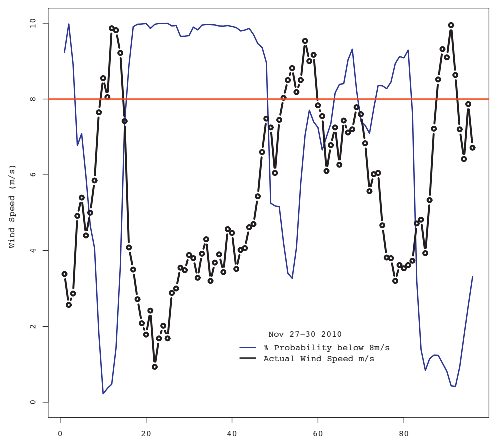

Fig. 2. BMA "Threshold Forecasting": probabilities of wind-speed lower than 8 m s^{-1} for the last 4 days of Nov. 2010 (blue curve). Verifying hourly observations are shown by the black curve. Probabilities are scaled by a factor of 10 to fit on the graph.

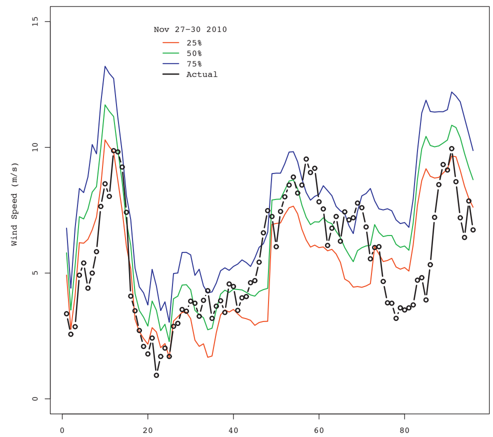

Fig. 3. BMA "Quartile Forecasting": 25% (red), 50% (green) and 75% (blue) chances of wind-speeds less than values shown. Verifying hourly observations are in black.

4.2. Forecasts with Operational HARMONIE for 2012

The operational forecast output produced 4 times daily by Met Éireann, using HARMONIE with a 2.5km resolution (corresponding closely to the "Arome" version mentioned above), was interpolated to the med-mast location and tested against the verifying wind-speed observations made there.

Since a new forecast run was started every 6 hours, each observed wind-speed could be used to validate 4 separate 24-hr forecasts (i.e., at 6, 12, 18 and 24hr lead-times). So 4 continuous forecast time-series were constructed from the 24hr forecasts started at 00Z, 06Z, 12Z and 18Z, respectively, each day. This constituted a small 4-member "ensemble" for the purposes of BMA analysis. Each member had lead times between 0 and 23 hours to the validating observation, and so over long enough times, each member should be statistically equivalent to each other.

Another way to construct an ensemble is, for each observation time, to let the first member be the forecast with the shortest lead time (i.e., 0 to 5 hours), the 2nd member be the forecast with lead times of 6-11 hours, the 3rd member the forecast with lead times of 12-17 hours, and the 4th member the forecast with lead times of 18-23 hours. These ensemble members are not equivalent: the member with the shortest lead-time would normally be expected to make the most accurate forecasts.

This is not necessarily the case, however, as shown in Table 2 for the period Jan.–Mar. 2012. The MAE results in Table 2 suggest that the first 6 hours of each forecast is typically a period of adjustment to new initial conditions and generates relatively large MAEs when validated against observations. The lead- time period with smallest errors is the 6–11hr interval, probably since by this time the forecast fields have adjusted to the "shock" of new initial and boundary conditions (however reduced by data assimilation methods), while the inevitable forecast errors have not had the opportunity to grow too large.

As also shown in Table 2, BMA post-processing (having been "trained" on 20 days of forecasts from January) does add value in this case, reducing overall MAEs by 9–16%.

Table 2 MAE for Jan.−Mar. 2012 from ensemble constructed from 4 separate forecasts, each with different lead-times. Also shown is BMA forecast (after 20 days "training" during Jan. 2012).

"Ensemble" member (with different lead-times to observation time)	MAE (m s^{-1})
0 – 5 hrs	1.80
6 – 11 hrs	1.71
12 – 17 hrs	1.78
18 – 23 hrs	1.79
BMA Forecast (from 4-member ensemble)	**1.55**

Time-series of wind-speeds forecast by the "6–11 hr" ensemble member (red), and by BMA (aqua), along with the verifying observations, are shown in Fig. 4 from this 65-day period in 2012. All the major observed oscillations are captured by the forecasts. There is a hint of a positive bias in the "raw" HARMONIE forecasts (red), but if so, this is precisely what BMA analysis (aqua) is designed to remove.

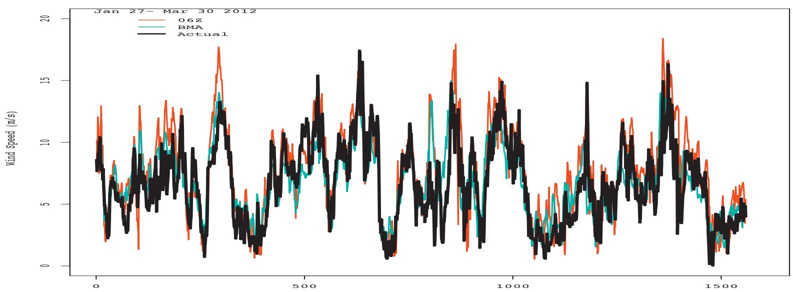

Fig. 4. Timeseries from 27 Jan. – 31 Mar. 2012 of observed winds (black), the most accurate "6 – 11hr" forecast ensemble member (red), and the BMA forecast (aqua), based on 4 ensemble members and 20 days training.

4.3. Forecasts with Operational HARMONIE for 2012

A 30-day suite of 24hr forecasts, starting 00Z and 12Z each day (with shorter 6hr forecasts in between for blending/assimilation purposes) was run from 10 Nov. 2012 through 9 Dec. 2012 with an experimental high-resolution version of HARMONIE. This version had 1 km horizontal grid-resolution and 65 vertical levels, compared with 2.5km horizontal resolution and 60 vertical levels in the operational HARMONIE. The computational domain of the high-resolution HARMONIE was about the same geographic size as that of the operational HARMONIE, so it had approximately a 2.5 x 2.5 or 6.25 times higher density of grid-points.

When measured against standard observations from weather observing stations, the high-resolution HARMONIE had overall slightly smaller biases and RMS errors than the current operational model. Of course, it was also a lot more computationally expensive to run.

At the wind-farm site, however, the high-resolution HARMONIE had typically too-strong winds, and an (uncorrected) MAE of 2.4 m s^{-1}. When this bias was removed, MAE was reduced to 1.84 m s^{-1}. No BMA post-processing was performed on the high-resolution output (there are only two "ensemble" members from this forecast suite to average in any case).

The full 30 days of wind-speed forecasts produced by this high-resolution HARMONIE at the met-mast, along with verifying observations, are shown in Fig. 5. As in all the other time-series shown above, the model captures the relatively low-frequency oscillations and slow variability quite well. The errors arise from over- or under-shooting of large-amplitude "spikes"; from real high-frequency (probably small-scale) events that are missed by the model, and from spurious high-frequency events that occur in the model.

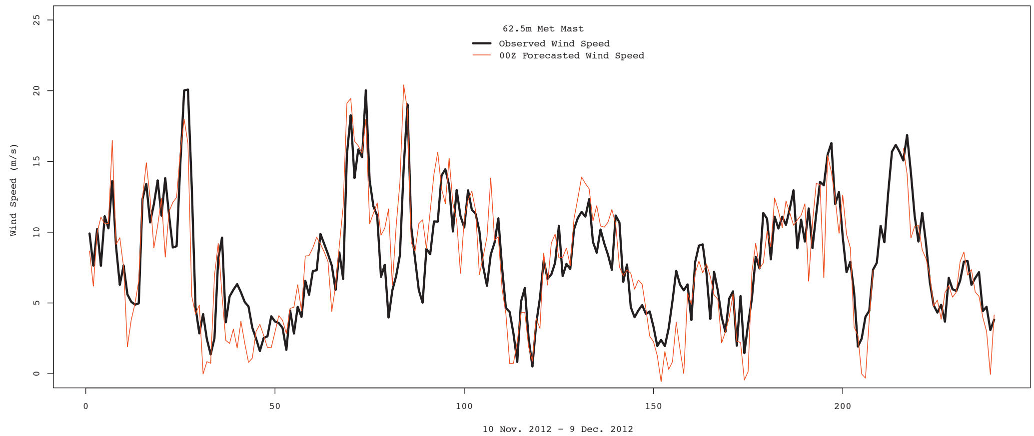

Figure 5 Timeseries from 10 Nov. – 9 Dec. 2012 of wind-speed at turbine height from the met-mast (black) and as forecast by the high-resolution HARMONIE (1km grid size) after local bias removal only. MAE is 1.84 m s^{-1}.

4.4. Site Dependency

Fig. 6 shows "raw" MAEs (i.e., unadjusted by BMA or other statistical post-processing) between standard 10m wind-speed forecasts from HARMONIE and validating observations, at a selection of weather stations around Ireland. MAEs are shown as a function of forecast lead-time (out to 24 hours), averaged over all of 2012. They were generated by the operational version of HARMONIE.

Clearly there is much variation between stations. The stations located in the (windier) west of Ireland tend to have larger MAEs, while those in the (less windy) east tend to have smaller MAEs. Clearly too, there is no uniform degradation in MAE as forecast lead-time increases. As shown in Table 2 for the wind-farm, forecasts are often most accurate at lead times of several hours rather than right at the start. Degradation of forecast accuracy is quite slow then thereafter.

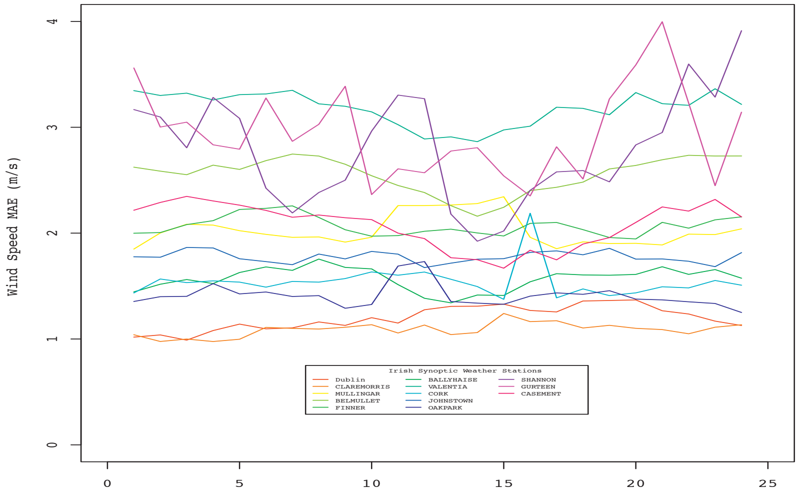

Figure 6 MAEs (m s^{-1}) between operational HARMONIE 10m wind-speed forecasts and observations at 13 weather stations around Ireland, as a function of forecast lead-time. Averages are over the full year 2012.

5. Conclusions

Operational HARMONIE forecasts generate MAEs of approximately 1.7 – 2.0 m s^{-1} when interpolated to the location and anemometer height of a wind-farm met-mast. Statistical post-processing (e.g., with BMA) can reduce this to approximately 1.5 – 1.6 m s^{-1}, for an improvement of about 10-15% overall.

MAEs do not degrade significantly as forecast lead-time increases, at least out to 24 hours. Indeed, the most accurate forecasts are typically not for the initial time, but rather have 6–11hr lead times.

A higher-resolution forecast model, which had slightly better skill-scores than the current operational model when validated against standard observing stations, nevertheless had a slightly higher MAE of 1.84 m s^{-1} over 30 days in late 2012 – though that was without the benefit of BMA post-processing.

Acknowledgements

We are grateful for financial support for this work from Irish EPA grant no. CCRP-09-FS-5-2. We also thank Met Éireann and Dr. Michael Sheehy of GaelForce Wind Energy for use of their data.

References

[1] Seity, Y, Brousseau P, Malardel S, Hello G, Bénard P, Bouttier F, Lac C, Masson V. The AROME-France Convective-Scale Operational Model. *Mon. Wea. Rev*. 2011, **139**, 976-991.

[2] Raftery AE, Gneiting T, Balabdaoui F, Polakowski M. Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *Mon. Wea. Rev*. 2005, **133**, 1155–1174.

[3] Gneiting T, Sloughter JM, Raftery AE. Probabilistic wind speed forecasting using ensembles and bayesian model averaging. *J. Amer. Statistical Assn*., 2010, **105**, 25-35.